

## Using Large Datasets of Organic Photovoltaic Performance Data to Elucidate Trends in Reliability Between 2009 and 2019

David, Tudur Wyn; Anizelli, Helder; Tyagi, Priyanka; Gray, Cameron; Teahan, William; Kettle, Jeffrey

**IEEE Journal of Photovoltaics**

DOI:

[10.1109/JPHOTOV.2019.2939070](https://doi.org/10.1109/JPHOTOV.2019.2939070)

Published: 01/11/2019

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

*Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):*

David, T. W., Anizelli, H., Tyagi, P., Gray, C., Teahan, W., & Kettle, J. (2019). Using Large Datasets of Organic Photovoltaic Performance Data to Elucidate Trends in Reliability Between 2009 and 2019. *IEEE Journal of Photovoltaics*, 9(6), 1768-1773.  
<https://doi.org/10.1109/JPHOTOV.2019.2939070>

### Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Using large datasets of Organic Photovoltaic performance data to elucidate trends in reliability between 2009-2019

Tudur Wyn David, Helder Anizelli, Priyanka Tyagi, Cameron Gray, William Teahan, *Jeff Kettle*

School of Computer Science and Electronic Engineering, Bangor University, Dean St, Bangor, Gwynedd, Wales, UK, LL57 1UT

**Abstract** — The application of data analytical approaches to understand long term stability trends of Organic Photovoltaics (OPVs) is presented. Nearly 1900 OPV data points have been catalogued and multivariate-analysis has been applied in order to identify patterns, produce models that quantitatively compare different internal and external stress factors and subsequently enable predictions of OPV stability to be achieved. Analysis of the weights associated with the acquired predictive model shows that for light stability (ISOS-L) testing, the most significant factor for increasing the time taken to reach 80% of the initial performance ( $T_{80}$ ) is the substrate and top electrode selection and the best light stability is achieved with a small molecule active layer. The weights for damp-heat (ISOS-D) testing shows that the type of encapsulation is the primary factor affecting the degradation to  $T_{80}$ . The use of data analytics and potentially machine learning can provide researchers in this area new insights into degradation patterns and emerging trends.

## I. INTRODUCTION

Stability remains a critical issue for researchers and industrialists in Organic Photovoltaics (OPVs) and Accelerated Life Testing (ALT) is regularly used, for example, to identify optimal material sets, predicting lifetime and providing relative comparisons of product stability [1]. To be successful, OPVs have to fulfil all three requirements of the ‘triangle’ – lifetime, efficiency and cost – simultaneously, otherwise they will be limited to a niche market.

OPV stability studies began in earnest towards the end of the last decade and have steadily increased as a result of its criticality upon the commercial potential of the technology and have been extensively reviewed [2,3]. Given the vast amount of data available on OPV stability, it might be possible to utilize large datasets containing results achieved across the community to better understand major patterns in OPV stability as well as enable lifetime prediction. Such a study could also improve the understanding of lifetime data as the conclusions of the experiments have been conducted with varying types and levels of stress factors including light, temperature and humidity. This is important, as the induced defects are likely to be affected by the inclusion of other stress factors and their ‘interactions’ [4]. We define ‘interaction’ as the relationship whereby the effect that a stress factor (e.g. light, temperature, humidity) has on the product is altered due to the presence of one or more other stress factors [5]. As an added level of complexity, stability is known to vary significantly dependent on material selection,

encapsulation strategy and number of thin film layers in a solar cell [24].

We propose using data obtained from the literature to identify the principle sources of degradation in solar cells when they have been tested with a range of environmental conditions and this data is used to identify stable material sets and device architectures. Nearly 1900 different samples have been catalogued for their structure and material composition with lifetime and performance data simultaneously recorded. A multivariate linear regression analysis (MLR) is employed in order to model the data and has been used to identify the most significant factors affecting the performance and lifetime of the devices. This method allows for the best configurations and combinations of attributes to maximize performance and stability to be identified. The data set acquired and analysed in this report contains 16 attributes with approximately 1900 instances. This leads to a dataset with a dimensionality of approximately 30,000. The total number of combinations which this dataset yields is vast and stands at ~972,000 combinations. Given such a large number of potential combinations, multivariate analysis of large datasets can be invaluable for sourcing trends in data.

## II. DATA ACQUISITION AND SCRUBBING

Data was obtained from two sources. Firstly, a data set was obtained from the DTU who ran the “lifetime predictor” on the Plastic photovoltaics website. Several parameters are included within this data set to describe 1) the structure of the solar cell, 2) it’s performance and 3) the stability under a wide range of test conditions. More recent data was extracted from publications using ‘Autodocsearch<sup>TM</sup>’, which is software that can be used for automated searching and organization of PDF documents. Only samples that had been tested in accordance with the ‘ISOS consensus testing protocols’ were selected in order to ensure consistency of the data [3]. In 2011, a consortium of researchers developed recommendations for standardized aging experiments which were known as the ‘ISOS protocols’. The vast majority of papers focus on using ISOS-L or ISOS-D protocols (these account for > 96% of reports), which report degradation under light or dark conditions, respectively. Dark storage studies (ISOS-D) provide information on the tolerance of OPVs to oxygen, moisture and thermal stress and therefore induce different failure mechanisms than those that occur during light soaking

(ISOS-L), which are related to photostability. Therefore, by assessing the performance under ISOS-L and ISOS-D separately, different degradation pathways can be investigated. In total, 1858 papers were used for the study.

The procedure for the data analysis approach used in this paper can be summarized by the OSEM process: Obtain, Scrub, Explore, Model, iNterpret [6]. “Obtain” and “Scrub” refer to the process of acquiring and cleansing the data. The “Obtain” stage is undertaken by acquiring data and ensuring it is in a suitable format such as the CSV, which can be read by a suitable database. To achieve this, lifetime metrics (discussed later) and data attributes (see Table I) had to be acquired from each paper. Attributes were chosen based upon previous literature as these had been shown to affect the OPV stability [6]. Scrubbing is performed in order to acquire a consistent format throughout the dataset. Each attribute is grouped by the categories listed in Table I. A schematic of the device configuration used for this study is shown in FIG.1.

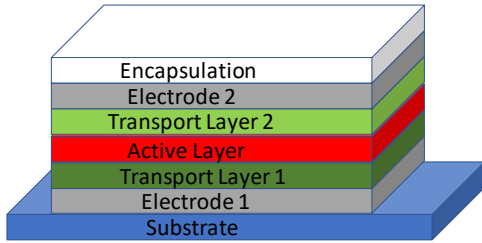


FIG.1: Structure of OPV devices and modules highlighting the categories defined for this work.

Attributes	Categories (Low to High values)
Configuration	Normal (L)   Inverted   Tandem (H)
Encapsulation	N/A   None   Flexible   Rigid
ISOS compatibility	L,D-1   L,D-2   D-3
Substrate	Glass   PET
First Electrode	ITO   Ag   Other Metal
First Transport Layer (TL1)	None   PEDOT:PSS   Low $\Phi$ Metal Oxide   High $\Phi$ Metal Oxide
Active Layer	Polymer   Small Molecule
Second Transport Layer (TL2)	Inorganic   Small Molecule   None
Second Electrode	Al   Ag   Au
Publication Year	2019 - 2009
Light Spectrum	N/A   Metal halide lamp   Sun   LED   Halogen lamp

TABLE I: Attributes and categories used in the dataset.

For each attribute, the lifetime metrics are acquired from all publications for consistent comparison against the attributes. The lifetime metrics selected were:

- Initial efficiency ( $E_0$ )
- Time to reach 80% of the initial efficiency ( $T_{80}$ )
- Time to reach 80% efficiency after the ‘burn in’ process has stopped ( $T_{S80}$ )

Finally, for analysis the data needed to be ‘summarized’ appropriately. This involves the categorization of similar instances into a single group. For example, all metal oxides are classed as ‘Low work function’ or ‘High work function’ Metal oxides, rather than  $\text{MoO}_3$ ,  $\text{V}_2\text{O}_5$  and  $\text{ZnO}$  etc. individually. This stage is needed to minimize the complexity of the data structure and, consequently, reduces the ‘over-fitting’ of data. Over-fitting is where any model becomes specialized to a particular set of data. The accuracy rate on this data set tends towards 100%, but at the same time drops dramatically on other data sets. By summarizing the data, with reduced complexity, the statistical significance of prominent trends can be identified and applied to the predictive MLR model.

### III. MULTIVARIATE LINEAR REGRESSION

A MLR model was made using a least square polynomial model to express the lifetime metrics in terms of the stress factors or environmental conditions. This was calculated by assuming this model of the characteristic lifetime as shown in eq. 1 [8]

$$y_i = f_0 + \sum_{k=1}^z f_k x_{nk} + \varepsilon_n \quad (1)$$

where  $y_i$  represents the response,  $f_0$  is a constant fitted parameter,  $f_1$  to  $f_k$  are the fitting parameters for the individual coefficients,  $x_{nk}$  represents the  $n^{\text{th}}$  level of the  $k^{\text{th}}$  predictor variable and  $\varepsilon_n$  represents the standard variance error. Multiple ‘y’ observations (representing the  $E_0$ ,  $T_{80}$  and  $T_{S80}$ ) can be expressed with the matrix in eq. 2.

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdot & x_{1n} \\ 1 & x_{21} & x_{22} & \cdot & x_{2n} \\ \vdots & \vdots & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \cdot & x_{nn} \end{bmatrix} \quad (2)$$

The matrix  $X$  is denoted as the design matrix containing information about the level of the predictor variable. In order to identify a regression line, the standard error must be accounted for. The standard error of the estimate is a measure of the accuracy of predictions and given in eq. 3.

$$\sigma_{est} = \sqrt{\frac{\sum (Y_{real} - Y_{est})^2}{N}}, \quad (3)$$

where  $\sigma_{est}$  is the standard error of the estimate,  $Y_{real}$  is an actual value and  $Y_{est}$  is an estimated value. The numerator is the sum of the squared differences between the reference values and the predicted values. The method of least squares is employed in MLR, whereby the coefficients in the matrix, shown in eq. (1) can be estimated by finding values of the coefficients that minimize the sum of the squared residuals [7].

In order to implement a MLR model, several criteria must be fulfilled. For all instances of the independent variable, the dependent variable should be normally distributed with

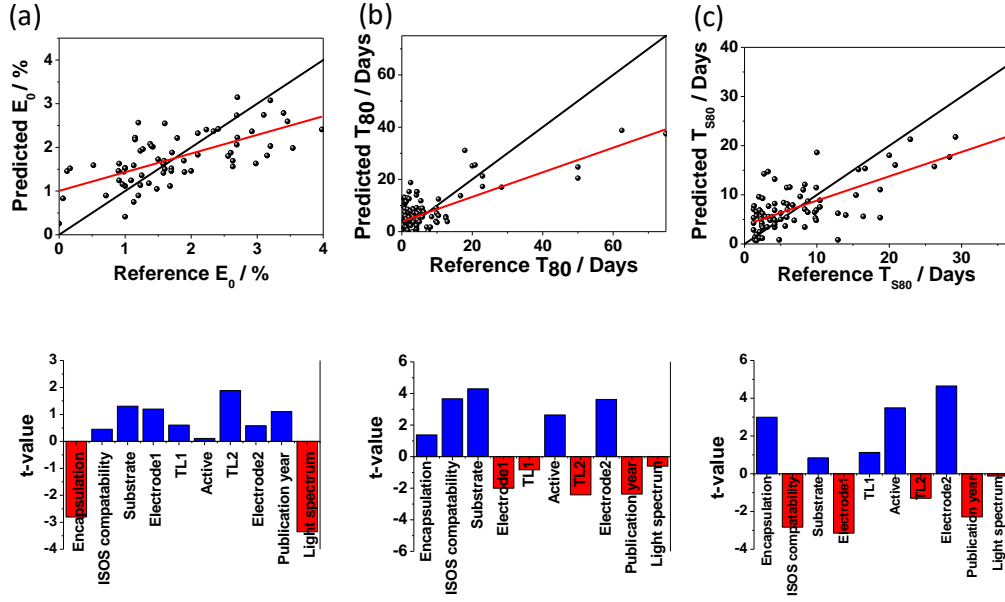


FIG.3: (a) Predicted vs. reference regression fit for  $E_0$  ISOS-L using MLR and associated  $t$ -values for each attribute used in MLR analysis. (b) Predicted vs. reference regression fit for  $T_{80}$  ISOS-L using MLR and associated  $t$ -values for each attribute used in MLR analysis. (c) Predicted vs. reference regression fit for  $T_{S80}$  ISOS-L using MLR and associated  $t$ -values for each attribute used in MLR analysis.

constant variance. In addition, the dependent variable and independent variable should scale linearly and all observations should be independent of one another (independence, linearity, normality, homoscedasticity). The efficacy of these criteria can be tested through residual analysis. If the criteria are fulfilled and the model is valid, then the residuals should be normally distributed with a mean of zero and a constant standard deviation. MLR is applied to the data set described in section II with the same attributes (factors) and responses in the MLR analysis. The  $t$ -value can be utilised as a metric assessing the significance of each factor in the prediction of either  $T_{80}$ ,  $T_{S80}$  or  $E_0$ . The  $t$ -value is defined as being,

$$t = \frac{\bar{x} - \mu}{SE}$$

Where  $\bar{x}$  is the mean of the sample,  $\mu$  is the mean of the population and  $SE$  is the standard error. The higher the  $t$ -value, the more significant the factor is in the prediction of the related response.

#### IV. RESULTS AND DISCUSSION

The MLR technique has been applied to the accelerated light-degradation studies (ISOS-L) and accelerated temperature/humidity studies (ISOS-D) separately. The criteria for verifying the validity of the MLR regression can be tested by analysis of the residuals between the predicted models and the reference data. Whilst regression data is not shown, the independence criteria are satisfied well for the ISOS-D data and

the  $E_0$  and  $T_{80}$  from the ISOS-D data although, some minor trends were observed in the  $T_{S80}$  data. The normality of the residuals holds very well since the predicted residuals are equally distributed above and below the 0-residual axis. Furthermore, the homoscedasticity criteria are fulfilled very well for the ISOS-D and the ISOS-L,  $E_0$  prediction, where the residuals are uniformly distributed across the space. However, for all  $T_{80}$  and  $T_{S80}$  data a clearer trend can be observed. However, if only the data with predicted lifetimes less than 45 days were considered, then all the criteria would be satisfied. This can explain the inaccuracies observed when predicting high  $T_{80}$  lifetimes. Finally, the linearity criteria are satisfied from the linear correlation between the predicted and reference values of the response parameters

FIG.3 shows the regression results obtained from papers restricted to light-degradation studies only, showing how the ‘Attributes’ affect the  $E_0$ ,  $T_{80}$  and  $T_{S80}$  values. In FIG.3(a), the results of the predicted vs. reference for the  $E_0$  calculation is shown calculated from the ISOS-L data. The desired feature of this plot is a fitted line with a slope close to one and all points lying close to the fitted line, indicating a close agreement between the regression model and the actual results. The correlation between the predicted and actual values can be represented by Pearson correlation ( $R^2$  value). For all data, the residual distribution plot can be found in the supplementary information (SI) section S1. However, FIG.3(a) shows an  $R^2$  value of 0.427 for the prediction of  $E_0$  using MLR. Given that this data is sourced from multiple authors in different groups, it is not surprising the ‘ideal’ case where the  $R^2$  value 1.0 is not observed, but FIG.3(a) provides some confidence that correlations can be found in the data set.

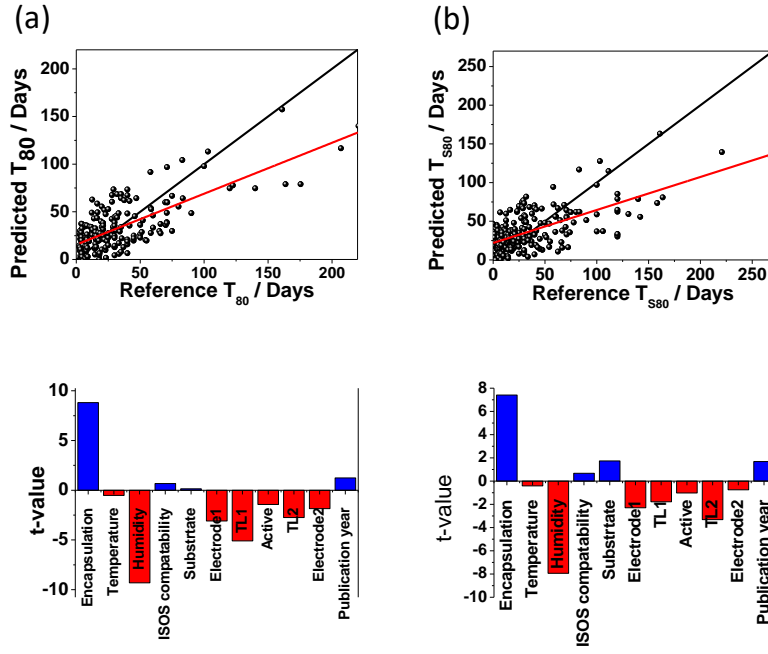


FIG.4: (a) Predicted vs. reference regression fit for  $T_{80}$  ISOS-D using MLR and associated  $t$ -values for each attribute used in MLR analysis. (b) Predicted vs. reference regression fit for  $T_{80}$  ISOS-D using MLR and associated  $t$ -values for each attribute used in MLR analysis.

Using MLR analysis, the regression  $t$ -values can be calculated from the individual regression coefficients of each attribute, divided by the standard error. A positive  $t$ -value indicates that the higher level of the attribute increases stability and the lower level reduces stability (with respect to the low and high values in Table I). Shown in FIG 3(a) are the  $t$ -values for  $E_0$  obtained from the ISOS-L. The regression  $t$ -value distribution indicates that the Encapsulation and Light spectrum are the attributes which have the most significant, negative impact on the  $E_0$  prediction whilst the most positive for improving the  $E_0$  prediction is the Transport Layer (TL) 2.

In the case of Encapsulation and Light spectrum, the results show that, publications with an undefined ‘N/A’ encapsulation and light spectrum give the highest efficiency as these are the low levels for their respective attributes in Table 1. The reason for this is actually a secondary effect; these data points correspond to small ‘cells’ rather than R2R manufactured modules. Generally, modules are always encapsulated, but they also tend to possess a lower efficiency than cells [8-9].

Whilst the  $E_0$  data provides some insight, the focus of the paper is on stability. FIG. 3 (b) shows the regression results for  $T_{80}$  values obtained only with ISOS-L data with the corresponding regression  $t$ -values shown in FIG 3 (b). The residual distribution plot can be found in the SI (section S2). FIG.3 (b) shows a good  $R^2$  value of 0.59 for the prediction of  $T_{80}$  using MLR, showing a moderately better fit than the  $E_0$  data. The regression  $t$ -values also allow us to study trends in OPV stability; the greater the  $t$ -value, the greater the impact this value has upon increasing the  $T_{80}$  metric.

In rank order, this indicates that ‘Substrate’, ‘ISOS compatibility’, ‘Electrode 2’, ‘Active layer’ and ‘Encapsulation’ type positively affects stability and that that ‘Transport Layer 2’, ‘Electrode 1’ and ‘Publication year’ negatively affect stability (in this rank order). Interestingly the substrate type  $t$ -value suggests PET is more stable than glass, as it shows a positive  $t$ -value. Module sized OPVs have been shown to possess greater stability in comparison to cells [8-12] and most modules are made on PET substrates (94% of instances in this dataset), which accounts for the greater stability. The data also suggest that gold electrodes are the most stable and that testing with ISOS-L-2 gives more stable devices than ISOS-L-1. The latter is peculiar as ISOS-L-2 is a more severe test. However, from analysis of the dataset distribution, 49% of devices tested under ISOS-L-1 were encapsulated, whereas 62% of devices tested under ISOS-L-2 were encapsulated, indicating that the interaction of the two factors is leading to greater predicted stability.

We have observed a contradictory pattern for  $T_{80}$  and  $T_{S80}$  in the  $t$ -value graph where ISOS compatibility shows a positive effect for  $T_{80}$  and a negative effect on  $T_{S80}$ . This may be attributed to the fact that 65% of the dataset possess P3HT:PCBM as the active layer. In P3HT:PCBM based OPVs, an improvement in performance when annealed due to an increase in carrier mobility and reduced trap density in the active region [12]. In case of ISOS-L-2, the devices are aged at a higher controlled temperature of 65/85°C. Therefore, during the ‘burn-in’ stage of the devices, the trap density within the active layer reduces, leading to higher  $T_{80}$  times. However, temperature in the long term can cause loss of conjugation in

polymers and can damage the contacts, which is reflected in the  $t$ -value graph for  $T_{S80}$  [13].

The data also indicates that overall for ‘Active layer,’ small molecules provide better stability than polymers, that inorganic transport layers (i.e. low function metal oxides) are more stable than other types of transport layers (TL1) and that ‘Publication year’ has an impact upon stability. Within our dataset, there are reports indicating that small molecules have very good stability to light soaking experiments due to their excellent morphological and photochemical stability [14,15]. However, 9.8% of papers in the dataset use small molecule active layers (in comparison to 88.9% polymer-based), so it is possible that a small number of high stability reports affect the conclusions. For the ‘Publication year,’ it appears that as time progresses, groups are enhancing PV stability which is not surprising as lessons are learnt from previous tests and thus stability increases.

FIG.3 (c) shows the predicted vs. reference and regression  $t$ -values regression analysis for  $T_{S80}$  ISOS-L data which yields the worst  $R^2$  value of 0.348. Whilst  $T_{S80}$  often quoted in OPV papers, it is possibly the most inaccurate to calculate because it requires the researcher to define when ‘burn-in’ has stopped, which can be subjective. Overall, the trends in the  $T_{S80}$  data follow the  $T_{80}$  data, although the ranking does change. There is just one exception; ISOS-L-1 is shown to produce better stability (as discussed earlier). For,  $T_{S80}$  the three most important factors are the top electrode, the active layer and the encapsulation. The increased significance of the active layer for the  $T_{S80}$  prediction could be associated with the fact that  $T_{S80}$  is the time to reach 80% efficiency after the burn-in process. This means that for  $T_{S80}$ , the trapped charge states in the active layer can now participate in charge extraction [13].

Interesting features can be observed when comparing FIG.3 (a) and FIG.3 (c). There exists a change in parity for the ‘Electrode 1’;  $E_0$  displays a positive correlation with the choice of Electrode 1 whilst  $T_{S80}$  displays a significant negative correlation with the choice of this electrode. This means that in order to optimize  $E_0$  and  $T_{S80}$ , a compromise must be reached which maximally satisfies both  $E_0$  and  $T_{S80}$ .

FIG.4(a) and (b) display the regression analysis for samples only subjected to temperature/humidity studies (based on ISOS-D data) for  $T_{80}$  and  $T_{S80}$  respectively. The analysis here is restricted to  $T_{80}$  and  $T_{S80}$  since analysis of  $E_0$  is not relevant when considering ISOS-D degradation. Humidity has been included in this analysis since many ISOS-D tests investigate the performance and stability of OPV devices under various humidity levels. FIG. 4 (a) displays a strong correlation in the prediction with an  $R^2$  value of 0.589 for  $T_{80}$  prediction and an  $R^2$  value of 0.517 for the  $T_{S80}$  prediction.

There are two primary attributes that contribute towards the ISOS-D  $T_{80}$  prediction: ‘Encapsulation’ and ‘Humidity’, which have a large positive and negative correlation, respectively. This is fairly obvious; testing with high humidity on non-encapsulated cells is obviously likely to rapidly degrade solar cells. Most other attributes have a negative impact on the  $T_{80}$  prediction except for ‘ISOS-compatibility’ and ‘Substrate’ which show a small positive

correlation. From the analysis, the stability of polymer-based OPVs is under greater stability thermal/humidity testing than small molecules; within the dataset there are few reports of highly stable small molecules under damp-heat conditions, in contrast to the polymer-based OPVs.

The next most impactful attributes resulting in a negative  $t$ -value (other than humidity and encapsulation) are TL1 and the Electrode indicating that using ITO and metal oxide provides the best stability. The regression analysis for  $T_{S80}$  using the dark data displays very similar trends as observed for  $T_{80}$ , which shows the same contributing factors. The effect of burn-in processes can also explain the differences between the ISOS-L and ISOS-D  $T_{80}$  and  $T_{S80}$   $t$ -values. It is noteworthy that there is a change in parity when comparing the ISOS-L and ISOS-D  $T_{80}$  and  $T_{S80}$   $t$ -values. For ISOS-L testing, the use of a small molecule active layer has a significant impact on enhancing the stability whilst for ISOS-D testing, using a polymer active layer improves the stability. This indicates that small molecule based active layers have better light stability and are better suited to higher irradiance condition. .

---

## V. CONCLUSION

A multivariate approach has been applied to analyze and understand the patterns and structure of a dataset containing OPV performance and stability information. The multivariate linear regression algorithm is adopted in order to analyze the dataset, thereby allowing  $E_0$ ,  $T_{80}$  and  $T_{S80}$  to be predicted for ISOS-L data and ISOS-D data. The quality of the regression fit can be quantified by the  $R^2$  metric. The regression  $t$ -values are also extracted which quantify how significant a particular parameter is for the model predictions, thus, allowing the most significant factors affecting the performance and stability to be ascertained.

To conclude from the ISOS-L and ISOS-D data, a number of general trends can be elucidated. For the ISOS-L data, as a general observation it is clear that the choice of spectrum plays a major influence on stability. Considering the design of the solar cell and both the  $T_{80}$  and  $T_{S80}$  ISOS-L data, the electrode selection appears to have the greatest influence on stability and that the use of metallic noble front and rear electrode yields the greatest stability. Active layer selection is also a major factor and more so that the choice the selection of first or second transport layer. Rigid encapsulation and glass substrates yield the best stability so for flexible OPV design, transferring the impermeable properties of glass to flexible substrates is vital. Considering the design of the solar cell and both the  $T_{80}$  and  $T_{S80}$  ISOS-D data, the major factor source of variation in stability data is the encapsulation technique, which is not surprising as this evaluates the tolerance of OPVs to oxygen and moisture. In contrast to ISOS-L data, the data indicates that the interface engineering has a more significant role in device stability and that the transport layers. Both ISOS-L and ISOS-D data show that as the publication year has an

influence on stability and as the progress in the field intensifies, stability is continuously improved.

One of the limitations of the Multivariate Linear Regression (MLR) approach is that we must group data into categories. Additional factors such as whether the active layer is planar or consists of a bulk-heterojunction, the composition of the active layer and the deposition technique can be included. However outliers of data can create significant noise in the model. Machine Learning (ML) algorithms might prove a better approach for analysing bigger datasets and can provide more powerful insights into the optimum structure of the OPV and produce more detailed predictions. ML algorithms allow for greater flexibility in the type of data analysed by employing more powerful regression techniques which subsequently allow for better predictive models by “learning” the behaviour and characteristics of the dataset. The application of ML and AI to future work could prove invaluable in predicting the future evolution of OPV devices, in particular for studying nuances such as the impact of solution processing, acceptor types etc.

## ACKNOWLEDGEMENTS

We would like to thank Dr Suren Gevorgyan for help with the initial data set. The authors would like to acknowledge funding by the European Regional Development Fund through the Solar Photovoltaic Academic Research Consortium (SPARC II) operated by the Welsh Government.

## REFERENCES

- [1] Schuller, S., P. Schilinsky, J. Hauch, and C. J. Brabec. "Determination of the degradation constant of bulk heterojunction solar cells by accelerated lifetime measurements." *Applied Physics A* 79, no. 1 (2004): 37-40.
- [2] Grossiord, N., Kroon, J. M., Andriessen, R., & Blom, P. W. (2012). Degradation mechanisms in organic photovoltaic devices. *Organic Electronics*, 13(3), 432-456.
- [3] Reese, M. et al. "Consensus stability testing protocols for organic photovoltaic materials and devices". *Solar Energy Materials & Solar Cells*. 95 (2011) 1253-1267.
- [4] Kettle, J., V. Stoichkov, D. Kumar, M. Corazza, S. A. Gevorgyan, and F. C. Krebs. "Using ISOS consensus test protocols for development of quantitative life test models in ageing of organic solar cells." *Solar Energy Materials and Solar Cells* 167 (2017): 53-59.
- [5] Stoichkov, V., Kumar, D., Tyagi, P. and Kettle, J., 2018. Multistress Testing of OPV Modules for Accurate Predictive Aging and Reliability Predictions. *IEEE Journal of Photovoltaics*, (99), pp.1-8.
- [6] Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. (2018). Big data with cognitive computing: a review for the future. *International Journal of Information Management*, 42, 78-89.
- [7] Esbensen, K. H. Swarbrick, B. Westad, F. Whitcomb, P. Anderson, M. Multivariate data analysis: an introduction to multivariate analysis, process analytical technology and quality by design. *CAMO Software AS*; 6th Edition, 2018
- [8] Jørgensen, Mikkel, Jon E. Carlé, Roar R. Søndergaard, Marie Lauritzen, Nikolaj A. Dagnæs-Hansen, Sedi L. Byskov, Thomas R. Andersen et al. "The state of organic solar cells—A meta analysis." *Solar Energy Materials and Solar Cells* 119 (2013): 84-93.
- [9] Bristow, N. and Kettle, J., 2018. Outdoor organic photovoltaic module characteristics: Benchmarking against other PV technologies for performance, calculation of Ross coefficient and outdoor stability monitoring. *Solar Energy Materials and Solar Cells*, 175, pp.52-59.
- [10] Angmo, Dechan, and Frederik C. Krebs. "Over 2 Years of Outdoor Operational and Storage Stability of ITO-Free, Fully Roll-to-Roll Fabricated Polymer Solar Cell Modules." *Energy Technology* 3, no. 7 (2015): 774-783.
- [11] Kettle, J., Bristow, N., Sweet, T. K., Jenkins, N., dos Reis Benatto, G. A., Jørgensen, M., & Krebs, F. C. (2015). Three dimensional corrugated organic photovoltaics for building integration; improving the efficiency, oblique angle and diffuse performance of solar cells. *Energy & Environmental Science*, 8(11), 3266-3273
- [12] Marsh, R.A., Hodgkiss, J.M., Albert-Seifried, S. and Friend, R.H., (2010) "Effect of annealing on P3HT: PCBM charge transfer and nanoscale morphology probed by ultrafast spectroscopy," *Nano letters*, 10(3), pp.923-930.
- [13] Zhong, C., Huang, F., Cao, Y., Moses, D. and Heeger, A.J., (2014), "Role of localized states on carrier transport in bulk heterojunction materials comprised of organic small molecule donors," *Advanced Materials*, 26(15), pp.2341-2345.
- [14] Burlingame, Q., Song, B., Ciannamarchi, L., Zanotti, G., Hankett, J., Chen, Z., Katz, E.A. and Forrest, S.R., 2016. Reliability of Small Molecule Organic Photovoltaics with Electron-Filtering Compound Buffer Layers. *Advanced Energy Materials*, 6(21), p.1601094.
- [15] Hermenau, M., Riede, M., Leo, K., Gevorgyan, S. A., Krebs, F. C., & Norrman, K. (2011). Water and oxygen induced degradation of small molecule organic solar cells. *Solar Energy Materials and Solar Cells*, 95(5), 1268-1277.